STUDENT ID NO

MULTIMEDIA UNIVERSITY

FINAL EXAMINATION

TRIMESTER 3, 2015/2016

TDM7011 – Advanced Data Management
(All sections / Groups)

16 MAY 2016
8.00pm – 10.00pm
(2 Hours)

INSTRUCTIONS TO STUDENTS

1. This question paper consists of 5 pages, including the cover page, with four questions only.

2. Attempt **ALL** questions in the paper. All questions carry equal marks and the distribution of the marks for each question is given.

3. Please print all your answers in the answer booklet provided.

QUESTION 1

Assume that you are working in a company that sells luxury cars which are imported from overseas. Your company is in the process of implementing a data warehouse for sales prediction.

(a) Extract, Transform and Load (ETL) are the processes of acquiring data from data sources and loading into the data warehouse.

You are asked to identify the **sources** for extracting relevant data for the data warehouse. What are suitable data sources? Briefly explain your answer.

[2 marks]

(b) You are asked to design the star schema for the data warehouse. Identify **four** dimensions and draw the star schema for the data warehouse, which will be mainly used for sales (fact) prediction purpose.

[5 marks]

(c) Data loading involves the processes of taking the prepared data, applying it to the data warehouse, and storing it in the database. There are four modes of applying data, which are load, append, destructive merge and constructive merge. Explain destructive merge and constructive merge using examples.

[3 marks]

Continued...

QUESTION 2

(a) Answer questions (i) to (iii) below based on the following table:

INVOICE

InvoiceNo	Date	Total	Discount	FinalTotal
R1001	2016-02-18	1000.00	5	NULL
R1002	2016-02-20	500.00	0	NULL
R1003	2016-02-20	1500.00	10	NULL
R1004	2016-03-22	1200.00	5	NULL

    (i)   Write the SQL code to create a trigger, which will set the value of *Discount* every time a new row is inserted based on the following:

Invoices with total of RM1000.00 and above will get 5% discount, and invoices with total of RM1500.00 and above will get 10% discount.

[3 marks]

    (ii)   Write the SQL code to update the *FinalTotal* of INVOICE.

[2 marks]

    (iii)   Write the SQL code which will return the following table when querying INVOICE table above.

Discount	No_of_Invoice
10	1
5	2

[2 marks]

(b) Due to the demand for higher storage, organizations have moved from a static fixed data center oriented operation to a more flexible and dynamic infrastructure for information processing, such as Storage Area Networks (SANs).

Explain SAN and give two advantages of SANs.

[3 marks]

Continued…

QUESTION 3

(a) State what are the 3 Vs related to data processing, and give an example of each V. Discuss how the challenges posed by the 3Vs can be solved using both a highly-distributed approach, as well as a high-performance approach.

[5 marks]

(b) Compare and contrast MapReduce and Hive by completing the following table.

[3 marks]

Criteria	MapReduce	Hive
Type of programming language		
Level of abstraction		
Number of lines of code		

(c) Discuss the advantages of Apache over MapReduce, from a developer's perspective.

[2 marks]

Continued...

QUESTION 4

(a) Map the following concepts in a relational database management system to their MongoDB equivalent:
- Table
- Row
- Column
- Relationship   [2 marks]

(b) You are provided with the following document located in a MongoDB instance.

```
{
"_id": ObjectId("5146bb52d85242700600001f3"),
"age": 25,
"city": "Los Angeles",
"email": "mark@abc.com",
"user_name": "Mark Hanks"
}
```

Convert the following entries in a relational database into the document format shown above, so that it can be imported into the same MongoDB instance.

[4 marks]

id	user_name	age	email	city
1	Tom Hanks	59	Tom_hanks@tomhanks.com	Los Angeles
2	Steven Gerrard	35	scouser@lfc.org	Los Angeles

(c) State and explain FOUR advantages of the XML database.   [4 Marks]

End of page.